13th International Conference on Applied Informatics
Eszterházy Károly Catholic University
Eger, Hungary, February 19–21, 2026

Bounded-Distance Prototype Reweighting for Noisy-Label Robustness in Image Classification

Mohammed A. Khudhair[a]

[a]University of Debrecen, Debrecen, Hungary
mohammeda.khudhair@inf.unideb.hu

Abstract

Label noise is a major obstacle for reliable medical image classification, where annotations may be inconsistent due to inter-observer variability and retrospective labeling. Our method will take a small *trusted* subset (assumed clean)that defines class prototypes in a normalized embedding space, and the remaining *untrusted* data are *softly* down-weighted based on their distance to the prototype of their (possibly corrupted) label. A key component is a *bounded* distance that saturates large coordinate-wise deviations, reducing the influence of outliers compared with unbounded metrics. On HAM10000 (7 classes) [3], with symmetric noise injected into untrusted labels at rates $\eta \in \{0, 0.3, 0.5\}$, prototype-Disk reweighting yields the following results: at (η=0.3), the robustness improved as validation accuracy went from 0.767 to 0.804 and macro-AUC from 0.920 to 0.939; at (η=0.5), accuracy improves from 0.740 to 0.789 and macro-AUC from 0.883 to 0.920 (mean over 3 seeds).

1 Introduction

Deep neural networks can memorize corrupted labels and degrade generalization, which is problematic in safety-critical domains such as dermatology. Existing methods either require repeated sample selection / co-training, or explicit noise modeling; here we ask a simpler question: can a small clean subset define a stable geometry that "softly rejects" inconsistent samples? Hence, we instead use a single trusted geometry + bounded distance to reduce noisy contributions softly.

2 Method: prototype disks with bounded distance

We use a ResNet50 [2] encoder followed by an ℓ_2-normalized embedding layer producing $\mathbf{z} \in \mathbb{R}^d$ (d=128) and a linear classifier. From trusted samples T_c of class c, we compute a prototype (mean embedding)

$$\boldsymbol{\mu}_c = \frac{1}{|T_c|} \sum_{(\mathbf{x},y)\in T_c,\, y=c} f_\theta(\mathbf{x}). \tag{1}$$

Bounded distance. For two embeddings, we measure deviation with a saturating, coordinate-wise distance

$$d(\mathbf{z}, \boldsymbol{\mu}) = \sum_{i=1}^{d} \frac{|z_i - \mu_i|}{|z_i - \mu_i| + \alpha}, \tag{2}$$

where $\alpha > 0$ controls saturation. Large deviations contribute at most 1 per dimension. Our bounded distance function limits the influence of outlier embeddings, drawing inspiration from robust distance metrics studied for noisy classification scenarios [1].

Prototype disks and reweighting. For each class we set a radius r_c as a high quantile (q=0.99) of $d(\mathbf{z}, \boldsymbol{\mu}_c)$ over trusted samples, defining a "disk" around $\boldsymbol{\mu}_c$. For an untrusted sample $(\mathbf{x}, \tilde{y})$ we compute $d_{\tilde{y}} = d(\mathbf{z}, \boldsymbol{\mu}_{\tilde{y}})$ and weight its cross-entropy by

$$w = \mathrm{clip}\Big(\sigma\big((r_{\tilde{y}} - d_{\tilde{y}})/T\big),\, w_{\min},\, 1\Big), \tag{3}$$

where T is a temperature and $w_{\min}$ avoids fully discarding samples. Intuitively, untrusted samples far outside their label's disk receive small weights and contribute less. We additionally use (i) a *purity* term encouraging confident predictions for samples inside disks and (ii) a *margin* term pushing trusted embeddings away from other-class disks.

Freezing the geometry. In our best-performing setting we freeze prototypes and radii during training (prototype update period set very large), avoiding prototype drift caused by noisy untrusted batches.

3 Experimental protocol

Split and noise. We use a stratified split (70% train / 15% val / 15% test). The training split is divided into trusted (15%) and untrusted (85%) subsets. Symmetric label noise is injected *only* into untrusted labels at rates $\eta \in \{0, 0.3, 0.5\}$ and repeated with seeds $\{1, 2, 3\}$. **Optimization.** We train for 8 epochs with AdamW, batch size 16 (CPU). We report validation accuracy and one-vs-rest macro-AUC. **Baseline choice.** The main baseline is the same ResNet50 trained directly on noisy labels using standard cross-entropy. This is a strong and fair baseline because it matches capacity and data processing, isolating the contribution of the proposed geometry and reweighting.

4 Results

Table 1 reports mean±std over three seeds, using the best validation AUC observed across the 8 epochs in each run. The baseline is strongest when labels are clean (η=0), where down-weighting is unnecessary. As noise increases, Prototype-Disk reweighting improves both accuracy and macro-AUC.

Table 1. Validation performance on HAM10000 under noise η. (Mean ± std over 3 runs).

Noise η	Baseline ResNet50		Prototype-Disk	
	Val Acc	Val AUC	Val Acc	Val AUC
0.0	0.8338 ± 0.0120	0.9655 ± 0.0043	0.8289 ± 0.0100	0.9526 ± 0.0066
0.3	0.7674 ± 0.0287	0.9201 ± 0.0132	0.8036 ± 0.0121	0.9388 ± 0.0097
0.5	0.7395 ± 0.0024	0.8827 ± 0.0260	0.7892 ± 0.0048	0.9197 ± 0.0061

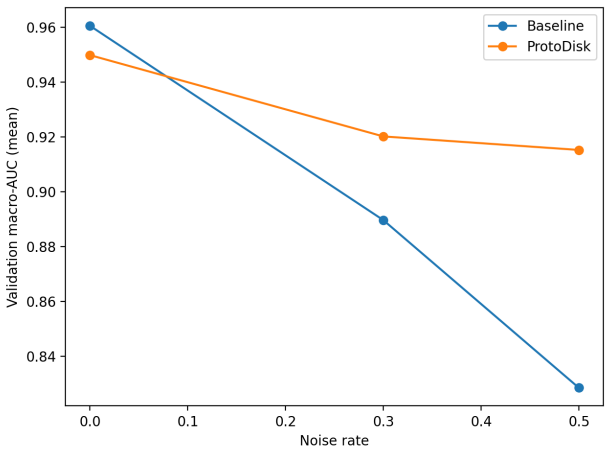

Figure 1. Prototype-Disk vs. Baseline Across Noise Rates.

In conclusion, Prototype-Disk reweighting with a bounded distance improves robustness as noise increases while keeping competitive clean-label accuracy. Future work will add stronger noisy-label baselines, study adaptive prototype updates, and validate on Real-World Noise with Multi-Label Data.

References

[1] H. A. Abu Alfeliat et al.: *Effects of distance measure choice on K-nearest neighbor classifier performance: A review*, Big Data 7.4 (2019), pp. 221–248, doi: 10.1089/big.2018.0175.

[2] K. He et al.: *Deep Residual Learning for Image Recognition*, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[3] P. Tschandl, C. Rosendahl, H. Kittler: *The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions*, Scientific Data 5 (2018), p. 180161, doi: 10.1038/sdata.2018.161.